

I Claim:

1. A voice activity detection apparatus, comprising:

- 5 a) an input for receiving an input signal derived from audio information, the input signal including a plurality of frames, each frame containing either one of active audio information and passive audio information;
- 10 b) a processing functional block coupled to said input for processing the input signal for generating an output signal capable to acquire at least two possible states, namely a first state and a second state, said first state being indicative of an input signal containing active audio information, said second state being indicative of an input signal containing passive audio information, said processing functional block being operative to:
- 15 i) for one or more frames received at said input and containing active audio information, compute a hangover time period, the computation including determining whether the hangover time period has a fixed duration or a variable duration, the determining being done on the basis of characteristics of the active audio information contained in the one or more frames;
- 20 ii) detecting a frame received at said input subsequently to the one or more frames containing the active audio information, that contains passive audio information; and
- 25

iii) causing the output signal to acquire said second state after the expiry of the computed hangover time period from the detecting of the frame containing the passive audio information.

- 5 2. A voice activity detection apparatus as defined in claim 1, wherein determining whether the hangover time period has a fixed duration or a variable duration is based on the duration of the active audio information contained in the one or more frames.
- 10 3. A voice activity detection apparatus as defined in claim 2, wherein if the duration of the active audio information contained in the one or more frames is less than a burst threshold, said hangover time period has a fixed duration.
- 15 4. A voice activity detection apparatus as defined in claim 3, wherein the fixed duration of said hangover time period is set to a predetermined constant value y .
- 20 5. A voice activity detection apparatus as defined in claim 3, wherein if the duration of the active audio information contained in the one or more frames is greater than the burst threshold, said hangover time period has a variable duration.
- 25 6. A voice activity detection apparatus as defined in claim 5, wherein the variable duration of said hangover time period is a function of the duration of the active audio information contained in the one or more frames.
7. A voice activity detection apparatus as defined in claim 6, wherein the one or more frames containing active audio information are characterised by a background noise energy level, whereby the variable duration of said hangover time

period is further a function of said background noise energy level.

8. A voice activity detection apparatus as defined in claim 1, wherein said processing functional block is operative to compute a classification data element for each frame of said input signal, the classification data element for a certain frame being indicative of whether the certain frame contains active audio information or passive audio information, a current state of the output signal being dependent at least in part on the basis of classification data elements computed with relation to previously received frames of the input signal.

9. A voice activity detection apparatus as defined in claim 8, wherein the classification data element is computed at least in part on the basis of a non-stationarity likelihood value associated with the certain frame.

10. A method for performing voice activity detection comprising:

a) receiving an input signal derived from audio information, the input signal including a plurality of frames, each frame containing either one of active audio information and passive audio information;

b) processing the input signal for generating an output signal capable to acquire at least two possible states, namely a first state and a second state, the first state being indicative of an input signal containing active audio information, the second state being indicative of an input signal containing passive audio information, the processing including:

- 5 i) for one or more frames received and containing
 active audio information, computing a hangover
 time period, the computing including
 determining whether the hangover time period
 has a fixed duration or a variable duration on
 the basis of characteristics of the active
 audio information contained in the one or more
 frames;
- 10 ii) detecting a frame received at said input
 subsequently to the one or more frames
 containing active audio information, that
 contains passive audio information; and
- 15 iii) causing the output signal to acquire the second
 state after the expiry of the computed hangover
 time period from the detecting of the frame
 containing passive audio information.
- 20 11. A method as defined in claim 10, wherein determining
 whether the hangover time period has a fixed duration or
 a variable duration is based on the duration of the active
 audio information contained in one or more frames.
12. A method as defined in claim 11, wherein if the duration
 of the active audio information contained in the one or
 more frames is less than a burst threshold, the hangover
 time period has a fixed duration.
- 25 13. A method as defined in claim 12, wherein the fixed
 duration of the hangover time period is set to a
 predetermined constant value y .
14. A method as defined in claim 12, wherein if the duration
 of the active audio information contained in the one or

more frames is greater than the burst threshold, the hangover time period has a variable duration.

15. A method as defined in claim 14, wherein the variable duration of the hangover time period is a function of the duration of the active audio information contained in the one or more frames.

16. A method as defined in claim 15, wherein the variable duration of the hangover time period is further a function of a background noise energy level in the one or more frames.

17. A voice activity detection apparatus, comprising:

a) input means for receiving an input signal derived from audio information, the input signal including a plurality of frames, each frame containing either one of active audio information and passive audio information;

b) processing means for processing the input signal for generating an output signal capable to acquire at least two possible states, namely a first state and a second state, said first state being indicative of an input signal containing active audio information, said second state being indicative of an input signal containing passive audio information, said processing means being operative to:

i) for one or more frames received at said input means and containing active audio information, compute a hangover time period, the computation including determining whether the hangover time period has a fixed duration or a variable

duration, the determining being done on the basis of characteristics of the active audio information contained in the one or more frames;

5 ii) detecting a frame received at said input means subsequently to the one or more frames containing the active audio information, that contains passive audio information; and

10 iii) causing the output signal to acquire said second state after the expiry of the computed hangover time period from the detecting of the frame containing the passive audio information.

14051ROUS02U